

Desarrollo de los ordenadores

1930. Teoría de la computación. Base de la teoría moderna de los calculadores modernos. Se demostraban cosas que podía hacer la máquina y cosas que no.

1940. Primera calculadora electromecánica, que ejecutaba operaciones automáticamente, codificadas previamente en cinta perforada en papel. Inventada en la Universidad de Harvard y se le llamó MARK-1.

1945. El ENAC, primera calculadora electrónica. Utilizaba tubos de vacío. Más rápida. Desarrollada en la Universidad de Pensilvania. Esta máquina no tenía almacenamiento de instrucciones, se introducían a través de un tablero de comunicaciones. Máquina especializada en balística.

1952. El mismo equipo diseñó una máquina que podía almacenar instrucciones y utilizaba números binarios para operaciones aritméticas.

Como continuación de estos prototipos universitarios surgen las primeras calculadoras de uso comercial. La primera la UNIVAC, que utilizaba cinta magnética y podía procesar datos alfanuméricos.

Todo este tipo de máquinas son consideradas como la primera generación.

1948. Aparece el transistor como sustituto del tubo de vacío. Más barato, más pequeño y apenas generaba calor. Revolución en la arquitectura de los aparatos electrónicos y sobre todo en los ordenadores. Comienzan a producirse en 1959, la segunda generación.

El paso de la segunda a la tercera generación se produce en 1964 con la aparición de los microcircuitos integrados de estado sólido (chips), que permiten más velocidad, más potencia de cálculo y más fiabilidad. Además desarrollo de software que permite uso más fácil del ordenador y desarrollo de sistemas operativos más potentes que permiten la multiprogramación y el multiproceso. Aparecen las familias de ordenadores (permiten realizar programas en una máquina y ser válidos para el resto de máquinas de la misma familia).

La cuarta generación. Aparecen las redes de ordenadores.

Etapas importantes en la arquitectura

1946. Tubo de vacío. Mucho más rápido que los relés.

1948. Transistor. 10 años de investigación para perfeccionar una producción masiva.

1946. Lógica del estado sólido. Técnicas para insertar diminutos circuitos en módulos de cerámica de media pulgada.

1970. Circuitos integrados. No se basan solo en transistores sino también en la tecnología monolítica a gran escala: compresión de los transistores y circuitos en una microplaqueta.

EL ORDENADOR

Componente principal de un proceso electrónico de datos.

Los ordenadores electrónicos utilizan componentes y circuitos electrónicos, en lugar de piezas mecánicas

movidas por motores eléctricos.

Tipos

El ciclo de las operaciones matemáticas puede realizarse de dos formas: analógica o digital. Los analógicos representan los números y cantidades así como los procesos de ciclo por variaciones de intensidad de corriente o variaciones de voltaje. Los digitales representan números por variaciones discretas de componentes físicos.

La velocidad de proceso en los ordenadores analógicos depende de la propagación de la corriente eléctrica mientras que los digitales dependen de la existencia de dispositivos electromagnéticos que disponen de límite de arranque, parada y movimiento.

Prácticamente hoy los ordenadores son de técnica digital aunque muchos son híbridos.

Componentes

Unidades de entrada, de salida, de control, Unidad Aritmético Lógica, memoria principal, memorias de almacenamiento externo.

El conjunto de la unidad de control, UAL y memoria principal es la Unidad Central de Proceso (CPU).

A medida que se van añadiendo cosas a un ordenador pasamos del microprocesador (CPU), a microordenador y microcomputador.

En la CPU no se pueden procesar los datos y producir información significativa sin la ayuda de otras unidades: de entrada, de salida, de almacenamiento de datos (periféricos).

- Microprocesador

Circuito integrado en el que reside la CPU.

En 1970 Intel desarrolla el primero. Diseño de un componente universal (funcionamiento no condicionado por el conexionado de sus componentes) nace el 4004.

• UAL: encargada de operar con los datos efectuando operaciones aritméticas y lógicas.

• Acumuladores: registro que almacena el resultado de las operaciones de la UAL.

• Contador del programa: indica la dirección de memoria.

• Registro de instrucción: almacena el código de la operación de la instrucción a ejecutar.

• Control: genera señales adecuadas para ejecutar la instrucción decodificada.

• Flags: señales que indican algún evento resultado de las operaciones de la UAL.

• Puntero de stack: registro que direcciona una zona de memoria destinada a contener direcciones de retorno de subrutinas.

Todos estos elementos están unidos por un bus interno. Hay tres tipos: bus de datos, de control y de instrucciones.

Todas las operaciones se realizan en tiempos preestablecidos cuya secuencia y duración las marca un reloj.

- Microcomputador

Conjunto formado por microprocesador, memoria RAM, memoria ROM y dispositivos de entrada y salida. Suele estar formado por una placa de circuitos impresos.

- Microordenador

El microcomputador más los periféricos. Los hay de entrada, salida y entrada/salida.

Estructura de un ordenador

- Flujo de datos: datos de entrada (valores iniciales), de proceso (resultados parciales) y de salida.

Los datos de entrada se alojan en la memoria principal y están a disposición de la UAL, donde se obtienen resultados parciales que se vuelven a almacenar en la memoria principal, volviendo a la UAL cuando sea necesario. Los datos finales quedan almacenados en la memoria principal, y finalmente son enviados a las unidades de salida.

- Flujo de instrucciones: las instrucciones se almacenan en la memoria principal y son enviadas a través de las unidades de entrada. Después se envían a la unidad de control donde se interpretan y se generan las órdenes para que se ejecute la instrucción.
- Flujo de control: la unidad de control por sí sola no ejecuta instrucciones sino que desde ella se emiten señales a las otras unidades para ejecutarlas.

Velocidad de proceso: suele medirse examinando la duración del ciclo de reloj de la CPU (tiempo recorrido para tener acceso a un carácter desde la memoria), más la velocidad de transformación de la información.

Velocidad de transferencia de datos: velocidad con que los datos entran y salen de la CPU desde o hacia los periféricos.

- Canales

Los periféricos se conectan a la CPU a través de ellos. Bajo el control del controlador de entrada y salida.

Un canal es un pequeño ordenador con capacidad para realizar sus propios programas llamados programas de canal, permitiendo al procesador desentenderse de las operaciones de entrada/salida.

Los canales pueden conectarse a varias unidades de entrada/salida.

Tipos: canales a rãfagas (selectores), donde el canal transmite toda la información debiendo las demás unidades esperar a que termine para comenzar la operación; canales múltiples (multiplexores), que transmite de cada vez un byte por cada unidad pasando los bytes por el canal entrelazados.

- Memoria principal: lugar de almacenamiento de la información.

La memoria se divide en departamentos denominados posiciones y cada una tiene asociado un número denominado dirección. Una posición únicamente puede almacenar un dato o una instrucción.

Dos tipos de operaciones: introducir o grabar la información (de forma autodestructiva, se sustituye la

antigua por la nueva) y extraer o recuperar la información (se puede efectuar la copia a otro dispositivo quedando la información inalterable en la memoria).

La cantidad máxima de información que puede tratarse en la memoria en una operación se denomina palabra. Dependiendo de cómo sean estas palabras nos da el tipo de memoria:

• Palabras de longitud fija. Fijan un número de bits que determinan la longitud de la palabra. Ordenador de palabra direccional.

• Palabras de longitud variable. La palabra tendrá un número variable de posiciones en función del dato almacenado. Se puede acceder a la información de dos formas: dando la dirección de la primera posición y un marca después de la última; o dando la dirección de la posición y el número de posiciones. Ordenadores de caracteres direccionales.

• Palabras direccionables a un objeto. Cada posición está formada por 8 bits. Una palabra está formada por 4 octetos y se accede a la información direccionando el octeto más a la izquierda.

El número de posiciones de la memoria principal mide la capacidad de la misma.

Estructuramiento físico: según la memoria de núcleos magnéticos, que ya no se fabrica pero contiene los principios básicos de las memorias actuales. Si atravesamos un anillo por un hilo conductor al que aplicamos una intensidad de corriente se crea alrededor del anillo un campo magnético dependiendo del flujo magnético de la intensidad de corriente aplicada. Además si el núcleo está compuesto por un material ferromagnético, este campo magnético es permanente. El sentido del campo magnético depende de la dirección de corriente que hayamos aplicado. Por tanto el núcleo puede estar en dos estados (cuerpo biestable) que se simbolizan por 0 y 1. Cada núcleo magnético va a ser un bit de información.

Cada núcleo está atravesado por dos hilos formando un ángulo recto. Al enviar a través de cada hilo la mitad de corriente necesaria para magnetizar un núcleo, sólo el núcleo que está en la intersección queda magnetizado. Para recuperar la información se enhebra un tercer hilo que se denomina sensor o conductor de detección. Se enhebra diagonalmente.

Estos núcleos son conocidos como núcleos de ferrita y se llaman memorias de ferrita.

- UAL

Donde se llevan a cabo las operaciones aritméticas, lógicas y de carácter intermedio; mover, desplazar y almacenar la información.

La unidades que permiten hacer esto son los sumadores, los registros, los comparadores y complementadores y los multiplexores y demultiplexores. Cada uno de ellos está formado por unos circuitos especializados. Los sumadores y complementadores están orientados hacia el cálculo, los comparadores hacia operaciones lógicas y los registros son utilizados indistintamente. Están basados en puertas lógicas.

• And o producto lógico. Equivale a circuito en serie

0	0	0
0	1	0
1	0	0
1	1	1

a-----

-----c

b-----

Â· Or o suma lÃ³gica. Equivale a circuito elÃ©ctrico en paralelo.

0	0	0
0	1	1
1	0	1
1	1	1

a-----

----c

b-----

Â· Complementador o negaciÃ³n.

a---- 0-----Ã

Â· Nand. NegaciÃ³n de la puerta and

0	0	1
0	1	1
0	0	1
1	1	0

a----

0----c

b-----

Â· Nor. NegaciÃ³n de la puerta or.

0	0	1
0	1	0
1	0	0
1	1	1

a----

0----c

b----

Â· Xor

0	0	0
---	---	---

0	1	1
1	0	1
1	1	0

a----

----c

b----

*Comparadores: esencia de los circuitos lógicos. Comparan parejas de bits de izquierda a derecha hasta que se determina cual es el más grande.

*Registros: conjunto de elementos biestables semejantes a las posiciones de memoria.

Tipos: de desplazamiento (desplaza la información hacia la derecha o la izquierda un cierto número de posiciones); acumuladores (almacenan los resultados de operaciones anteriores); multiplicador y divisor (almacena el multiplicador de un producto o el divisor de una división); generales (circuitos capaces de recibir, conservar y transmitir información con las órdenes de la UAL)

Se pueden producir dos tipos de situaciones: overflow (el valor del dato a almacenar excede las dimensiones de un registro) y underflow (el valor a almacenar en un registro es tan pequeño que no se detecta).

- Unidad de control

Dirige y coordina todo el proceso de datos. Depende de las órdenes e instrucciones de nuestro programa.

Funciones:

1º Recoger las instrucciones del programa almacenado en la memoria principal, de acuerdo con la secuencia fijada por el programador

2º Interpretar las instrucciones a ejecutar

3º Establecer las conexiones entre las partes de la máquina de acuerdo con la interpretación anterior.

Elementos:

• Registros: la misma naturaleza que los de la UAL. Hay registros de almacenamiento (almacena el operador), de instrucción (almacena el operando), y de dirección (almacena la dirección de memoria de los operandos).

• Codificadores y Decodificadores: el codificador convierte los datos codificados en decimal a binario. El decodificador al revés.

• Contadores (identificadores de instrucción): siempre tienen almacenada la dirección de memoria de la instrucción que tiene que ejecutar.

• Reloj: todas las operaciones tienen que realizarse en un tiempo y secuencia preestablecida. Utilizan unas señales denominadas impulsos de reloj, que son enviados a todos los componentes para sincronizar.

Almacenamiento externo

- Los discos

Su funcionamiento se basa en una propiedad física de sus componentes por la que unas partículas quedan orientadas al aplicarles un campo magnético. Dependiendo del sentido de orientación, asumimos un 0 o un 1, almacenando los bits de información. Para leer la información se utiliza un cabezal que detecta la orientación de las partículas.

• Soportes

Por la forma de almacenar los datos: magnética, óptica o magnetoóptica.

Por su movilidad e ubicación: internos o fijos; externos o removibles.

Por su modo de acceder a los datos: de acceso secuencial (hay que leer toda la información precedente); de acceso aleatorio o acceso directo (una vez conocida la ubicación de una información se accede directamente a ella). Todos los dispositivos que vamos a tratar a excepción de las cintas DAT son de acceso aleatorio.

El soporte fijo más utilizado es el **disco duro**. Suele estar en el interior del ordenador aunque hay modelos extraíbles.

Se suele comunicar con el procesador a través de distintos tipos de buses, como son el IDE, el EIDE y SCSI.

Puede tener una (sistemas de brazo o cabeza móvil) o varias cabezas lectoras por superficie (cabeza fija, una cabeza para cada pista)

Una pista es la parte del soporte que gira delante de la cabeza, una corona circular cuyos datos pueden ser leídos por la cabeza.

A su vez las pistas se dividen en sectores. Un sector es la unidad de información que se transfiere en un acceso. Su forma geométrica es la de un sector circular. Todas las pistas tienen el mismo número de sectores, pero son de distinto tamaño. Aumenta a medida que nos alejamos del eje.

En los dispositivos con una superficie, un cilindro está constituido por las pistas concéntricas.

El disco tarda un tiempo en alcanzar una velocidad estable. Dicha velocidad se denomina velocidad nominal de rotación, a la que se puede leer o escribir datos. El disco suele estar siempre girando, para no perder tiempo en cada operación.

Se calibran las cabezas lectoras para que pasen a una distancia mínima sobre la superficie del disco, gracias al aire que desplaza el disco al girar. Hace que la cabeza planee, adaptándose incluso a las irregularidades de la superficie.

En los dispositivos extraíbles las cabezas se apoyan sobre la superficie del soporte, y hace que su vida sea menor. Para alargar su vida el motor está por lo general parado, y solo se pone en marcha cuando hace una operación de lectura o escritura. Pero lleva a un incremento sustancial del tiempo de acceso.

El tiempo de acceso es el que tarda en posicionarse una cabeza en la pista deseada (tiempo de búsqueda), más el tiempo que tarda la información solicitada en estar justo debajo de la cabeza lectora (latencia). Estos tiempos no son siempre iguales ya que depende de la posición de partida como de la posición de destino. En las especificaciones técnicas se indica un tiempo promedio.

La velocidad de transferencia es la velocidad a la que se transfieren los bits de un sector, y viene determinada por la velocidad de giro del disco y por la densidad de los datos grabados en él. También hay que tener en cuenta el tipo de interface con el que se conecta al procesador, que sea capaz de absorber y transmitir todo el flujo de datos.

Superficie (2πr) X densidad de grabación X velocidad de giro = velocidad de transferencia

Interface SCSI

Tasa de transferencia de datos muy alta, y su principal virtud es que se mantiene constante en todo momento, sin que el procesador realice ningún trabajo. De gran importancia en procesos largos y complejos.

Son más caros que su equivalente IDE. Precisan de una controladora SCSI para manejarlos.

Permite conectar de 7 a 15 dispositivos frente al máximo de 4 de los IDE.

• Estructura del disco

No todos los sectores tienen la misma función, sino que dependen de su posición y del sistema de archivos que implemente el sistema operativo.

Hay un sector especial: el Master Boot Record (MBR) o registro de arranque principal, que está en el primer sector del disco, en el sector 0. Es el encargado del arranque del sistema operativo.

El resto de los sectores se dividen en: Área de datos (almacena los datos) y Área de localización (información que permite conocer qué sectores del Área de datos pertenecen a un determinado archivo, y cuál es su localización). En el entorno Microsoft esta Área se llama FAT (File Allocation Table).

Discos extraíbles

Permiten extraer, almacenar y guardar información fuera del ordenador o intercambiarla con otros ordenadores. Gran variedad.

• Disquette: medidas de 8 pulgadas, 5 1/4 y 3 1/2. Los de 3 1/2 pueden almacenar 720k, 1.44M o 2.88M (sólo en soportes IBM). Bajo coste, baja tasa de transferencia y tiempo de acceso elevado. Ventaja de tamaño manejable, su bajo coste y su estandarización.

• Discos ZIP: capacidad de 100Mb; similares a los disquetes de 3 1/2 aunque ligeramente más grandes. Tasa de transferencia y precio más altos que los disquetes. Se conectan al ordenador a través de unas unidades lectoras-grabadoras mediante puerto paralelo, puerto USB o interfaz SCSI. Aunque bastante extendidos no están estandarizados. Precio bastante superior comparándolo con los CDRW.

• Superdisco LS 120: 120Mb de capacidad. Distribuido por Panasonic. Mantiene compatibilidad con los disquetes de 3 1/2, y es más pequeño que un ZIP.

• Discos EZFlyer: similares a los discos ZIP, tanto en tamaño como en funcionamiento. Capacidad de 230Mb.

• Discos JAZ: capacidad de 1 a 2 Gb. Alta velocidad de transferencia. Elevado precio.

• Discos CDRW: no son magnéticos, son ópticos. Discos CD-ROM capaces de ser escritos múltiples veces. Capacidades entre 650 y 700 Mb. Precio reducido, pero reducida fiabilidad.

• Discos DVD: también ópticos. Gran capacidad de almacenamiento, precio asequible y compatibilidad con los CD-ROM.

• Cintas DAT: dispositivos magnéticos secuenciales (para acceder a un dato la cabeza lectora debe pasar por los que le preceden). Su principal uso las copias de seguridad. Gran capacidad, algo lentas pero buena relación calidad precio.

Sistemas de archivo: determina la estructura, nombre, forma de acceso, uso y protección de los archivos que se guardan en el disco. Cada sistema operativo tiene su propio sistema de archivo.

Nombre de archivo: las reglas varían de un sistema a otro, pero en general todos permiten cadenas de hasta 8 caracteres alfanuméricos, y utilizan nombres con dos partes separadas por un punto. A la segunda parte se le llama extensión e identifica el tipo de archivo que es. Suelen ser 3 caracteres alfanuméricos.

Atributos:

• s. Atributo del sistema (indica si el archivo pertenece al sistema operativo o no)

• h. Oculto (indica si el archivo está oculto)

• r. solo lectura

• a. Atributo de archivo (suele cambiar cuando se modifica un archivo. Permite determinar qué archivos se han modificado desde la última copia de seguridad)

• de fecha. Guarda la fecha de creación o modificación.

• de hora. Guarda la hora de creación o modificación.

• tamaño. Tamaño que ocupa el archivo.

Aparte de estos, en función del sistema operativo, hay atributos que indican la pertenencia del archivo a un usuario o grupo, y los permisos que el propietario tiene sobre él (lectura, escritura, ejecución), y el tipo de archivo que es (normal, enlace físico, enlace lógico, directorio).

En un sistema de archivos hay dos tipos principales de objetos: archivos y directorios. Los archivos son los encargados de contener los datos. Los directorios permiten una mayor organización de los archivos en el disco. Es un contenedor, puede contener archivos u otros directorios. Se puede llegar a crear una jerarquía en forma de árbol. Tiene la información necesaria sobre la posición del sector del disco en el que comienza, permitiendo al sistema operativo recuperar su contenido del disco.

Hay un directorio especial, el directorio raíz, que contiene todos los demás directorios y archivos. Es desde el que se parte cuando se busca un archivo mediante una ruta de acceso absoluto. Cuando se usa una ruta de acceso relativa el archivo se busca partiendo del directorio en el que se está trabajando, del directorio activo.

En MSDOS las rutas de acceso están separadas por \, mientras que en UNIX por /. Las rutas relativas carecen de carácter inicial. Muchos sistemas operativos tienen dos entradas especiales: . y .., que hacen referencia al directorio activo y a su padre respectivamente.

Un bloque está compuesto por un determinado número de sectores, que se asocian en un único archivo. Un archivo por tanto se almacena en uno o más bloques.

Elección del tamaño del bloque. Si es grande y el archivo pequeño se le asigna el bloque entero igual, desperdiciando gran parte del tamaño en disco. Si sucede lo contrario, harán falta muchos bloques produciendo un retraso en la lectura del archivo. Estudios indican que el tamaño medio de los archivos en sistemas UNIX y MSDOS rondan el kb, así pues son adecuados los tamaños de bloque de 512 bytes, de 1kb y 2kb.

Para manejar los bloques asociados a un archivo hay distintas técnicas:

• Asignación adyacente: almacenar los bloques adyacentes en el disco. Fácil implementación, pero hay que conocer con anterioridad el número de bloques que ocupará el archivo, y esto por lo general no ocurre. Genera una gran fragmentación del disco, lo que produce una pérdida de espacio.

• Asignación en forma de lista ligada: el directorio contiene la dirección del primer bloque, y cada bloque contiene la dirección del siguiente hasta el valor nulo del último bloque. Se evita la fragmentación del disco, pero cada bloque parte de su capacidad al tener que reservar espacio para la dirección del siguiente bloque. Produce retardo al tener que leer todos los bloques antes de llegar a uno determinado.

• Asignación mediante lista ligada y un índice: tabla con un registro por cada uno de los bloques que tiene el disco. En cada registro se indica si el bloque está vacío o ocupado. En caso de que esté ocupado se guarda la dirección del siguiente bloque. El directorio se asocia con el nombre del archivo el número de bloque en el que comienza y con la tabla se averigua la dirección del resto. El acceso a un determinado bloque es más rápido al estar la tabla residente en memoria, no tiene que acceder al disco, pero la memoria consumida no está disponible para otros procesos. Esto puede ser un problema en discos con un gran número de bloques, ya que la tabla de registros puede ocupar una gran parte de la memoria o incluso desbordarla. Es la técnica usada por el MSDOS (Microsoft), llamándose la tabla de registros FAT, pudiendo ser FAT16 o FAT32, dependiendo si los bloques se direccionan con 16 bits o con 32.

• I-Nodo: sistemas operativos como el UNIX. A cada archivo se asocia una pequeña tabla llamada i-nodo, que contiene los atributos y direcciones de los bloques del archivo. Las últimas entradas del i-nodo se reservan para cuando el archivo ocupa más bloques de los que el i-nodo es capaz de almacenar, y puede contener de esta manera la dirección de otro bloque en el que se guardan las direcciones de los demás bloques de archivo. A este bloque se le llama bloque indirecto. Si aún así no hay suficiente espacio para guardar las direcciones de los bloques del archivo se puede utilizar un bloque doblemente indirecto o incluso triplemente indirecto. Cuando se abre un archivo, lo primero que se hace es cargar su i-nodo en memoria para que el proceso sea lo más rápido posible.

Utilidades de desarrollo

- Programas de edición: facilitan la modificación de ficheros. Cada tipo de fichero tendrá su propio programa de edición, aunque los hay que editan múltiples formatos. Entre los más comunes están los editores de texto.
- Programas compiladores y enlazadores: permiten traducir el código fuente de un programa desarrollado a alto nivel en un determinado lenguaje de programación, a un código de bajo nivel entendible y ejecutable por el ordenador, al lenguaje máquina.

Cada compilador es específico para cada lenguaje y sistema operativo.

- Programas gestores de datos: gestor de bases de datos, generador de informes, plantillas para añadir o editar los datos. La base de datos es una estructura organizada de la información.

Tiene una estructura básica común: campos, que es cada una de las partes que componen el registro. Una tabla es una agrupación de registros. Y existen también las relaciones, que permiten crear estructuras

complejas relacionando varias tablas entre sí. Hay un campo especial, llamado clave principal, clave única o PK, que ha de ser único en cada tabla, e identifica de forma única a cada registro.

Uno de los mayores problemas que presentan las bases de datos es el tiempo empleado en recuperar la información, debido a que las bases de datos tienen grandes cantidades de datos.

Para ello se pueden utilizar las tablas indexadas, tablas que contienen un campo que es especial para cada registro, único y que además los registros están ordenados por él.

Una técnica más avanzada es el uso de tablas HASH, que funciona a través de claves. Cada llave es asignada o mapeada mediante una función a un número entero, en el rango entre 0 y $H_SIZE - 1$, donde SIZE es el tamaño de la tabla y 1 el valor que se le resta. Este número indicará la celda donde debe residir la información. Al mapeado se le llama función de HASH, que normalmente es más elaborada.

Sistema operativo

Un sistema operativo es un conjunto de programas que se encargan de gestionar y asignar recursos de hardware a los distintos procesos en ejecución.

Un proceso es un programa que está en ejecución con su entorno.

El sistema operativo colabora para que el ordenador se utilice de forma eficiente, tiene que controlar que la velocidad de los periféricos sea menor que la de la CPU, que los procesos de E/S mantengan ocupada a la CPU y debe controlar el acceso directo a memoria (DMA)

Evolución:

- 1950. Tarjetas perforadas. El sistema operativo controla y secuencia programas incorporados en tarjetas perforadas.
- 1960. Equilibrar la diferencia entre la velocidad de los periféricos y la CPU. Dos nuevos conceptos: procesos online (conectar los dispositivos directamente al ordenador) y procesos offline (conectar dispositivos lentos a través de dispositivos más rápidos). Aparte el buffering (utilizar memorias internas para el almacenamiento de datos) y el spooling (utilizar el disco como memoria interna).
- 1970. Aparece la multiprogramación: ejecución simultánea de varios programas por un solo procesador, y con ello surgen los procesos batch o por lotes (ir solicitando la ejecución de procesos que no necesitan la interacción del usuario, colocándolos en una cola de trabajo controlada por el sistema operativo).

Aparece también el tiempo compartido, un conjunto de terminales que están simultáneamente requiriendo la atención del microprocesador, siendo el sistema operativo el encargado de ir intercalando tiempos de usuario..

- Desde los 80. Se centran más en aumentar la seguridad de los sistemas operativos que en darles velocidad. Aparece el concepto de sistema de red. Como mejora de la velocidad aparece el multiproceso (el sistema operativo controla el reparto de trabajo entre los diferentes procesadores que componen el ordenador).

Arquitectura

- Monolítica. No hay estructura. Un conjunto de procedimientos que pueden llamarse entre sí, sin ningún tipo de limitación.

Para la construcción del sistema operativo completo se compilan de forma individual los archivos que contienen los procedimientos, y estos archivos se enlazan entre sí en un macroarchivo que es el sistema operativo.

- En capas.

• Capa 0: capa de abstracción del hardware; aisla el resto de capas de las particularidades que tiene el hardware.

• Capa 1: planificación y uso de la CPU; asignación del procesador, alternando entre los diversos procesos cuando ocurren interrupciones o se acaba el tiempo.

• Capa 2: administración de la memoria; asignar la memoria principal a los procesos y gestionar aquellos que no caben en memoria.

• Capa 3: gestiones de E/S; guardando en almacenes temporales o buffers los flujos de información que existen entre los dispositivos.

• Capa 4: interface de llamadas al sistema; procesos de usuario que requieren recibir servicios disponibles.

• Capa 5: programas de usuario; ejecuta los programas de los usuarios.

Cuanto más se acercan al hardware mayor prioridad de ejecución.

Los sistemas operativos evolucionan con el fin de desplazar la mayor cantidad de código a las capas cercanas al usuario, con la idea de mantener el núcleo con un tamaño mínimo (microkernel).

Interface de usuario: interface de programación de aplicaciones (API, DLL)

Un API es un conjunto de bibliotecas de programación, que elaboran y publican los fabricantes de sistemas operativos. En Windows está el kernel32.

Hay dos tipos de interfaz de usuario: modo texto, que se basa en un intérprete de comandos que marca las reglas sintácticas y semánticas; el modo gráfico (GUI), que va por encima del modo texto, haciendo que el sistema operativo sea más intuitivo y cómodo para el usuario.

Modos de explotación

- Según el número simultáneo de usuarios que soporte un sistema operativo: monousuario y multiusuario
- Según el número de procesos simultáneos: monotarea o monoprogramación (hasta que se acabe el proceso no se inicia el siguiente) y multitarea o multiprogramación (aunque en un momento dado la CPU sólo puede atender a un solo proceso).
- Según el número de procesadores: monoprocesador y multiprocesador
- Según los requerimientos de tiempo: en tiempo real (respuesta del sistema debe tener un intervalo de tiempo fijo), sistemas interactivos (depende de la velocidad de respuesta del usuario) y sistemas batch o por lotes (un proceso solo empieza a ejecutarse cuando el último proceso en posesión de la CPU termina).

Gestión de procesos

Un proceso es un programa en ejecución, que necesita un recurso del sistema para llevar a cabo una tarea.

Cuando se inicia un proceso se crea su bloque central de proceso, que carga la identificación del proceso, su estado (listo o bloqueado), el estado del hardware, su prioridad, los recursos solicitados y los asignados, la información sobre gestión de memoria y la información sobre los archivos.

- Planificadores: módulos del sistema operativo que se encargan de gestionar la ejecución de procesos. Para optimizar el rendimiento del sistema, usar al máximo la CPU (mayor número de procesos por unidad de tiempo)

• De admisión o de trabajo: seleccionar qué trabajo de los que se encuentra en disco pasa a la cola de los procesos listos en la memoria principal.

• Intermedio: llevar al disco los procesos de memoria (por agotamiento de la memoria principal o para establecer una buena combinación de procesos)

• De bajo nivel o de CPU: seleccionar entre los procesos listos los que van a ser atendidos por la CPU.

Los procesos deben sincronizarse, y hay dos tipos de sincronización: intercambio de señales entre procesos y la memoria compartida.

Para garantizar la integridad del sistema se utiliza la técnica de la exclusión mutua: cuando un proceso ejecuta su sección crítica debe completar todas las instrucciones de la sección antes de que otro proceso entre en la misma sección.

- Interbloqueos (abrazo mortal): cuando un grupo de procesos está bloqueando permanentemente un conjunto de recursos.

Deben cumplirse cuatro condiciones: que sea un recurso dedicado, retener y esperar, que exista apropiación y espera circular.

Estrategias: prevención (definir unas normas que aseguren que eso no ocurra), detección (el sistema operativo ejecuta una serie de algoritmos para saber si hay bloqueos, y si los hay detiene los procesos implicados) y la evitación.